Learning Pairwise SVM on Deep Features for Ear Recognition

Ibrahim Omara[1,2], Xiaohe Wu[1], Hongzhi Zhang[1], Yong Du[3] and Wangmeng Zuo[1]
[1]School of Computer Science and Technology,
Harbin Institute of Technology, Harbin, China 150001
[2]Department of Mathematics, Faculty of Science
Menoufia University, Shebin El-koom, Egypt 32511
[3]Department of Electrical and Information Engineering,
Northeast Agricultural University, Harbin, China 150001
Email: I_ omara84@yahoo.com, {xiaohe612, zhanghz0451, cswmzuo}@gmail.com, duyonghit@126.com

***Abstract*—Recently, deep features extracted from Convolutional Neural Networks (CNNs) have been widely adopted in various applications, such as face recognition. Compared with the handcrafted descriptors, deep features have more powerful representation ability which can lead to better performance. Effective feature representations play an important role in ear recognition. While deep features have not been applied to represent the ear images. In this paper, we propose to extract deep features of ear images based on VGG-M Net for solving the ear recognition problem. And due to the lack of training images per person, we propose to use the pairwise SVM for classification firstly. For computational efficiency, Principal Component Analysis (PCA) is exploited to reduce the dimension before classification. Finally, we evaluate our approach on two public ear databases: USTB I and USTB II. The experimental results achieve a promising recognition rate and show superior performance compared with the state-of-the-art methods.**

***keywords*—Ear recognition, deep features, CNN, pairwise SVM.**

I. Introduction

Biometric systems have drawn much attention in the public security domains, for providing personal identification or verification based on the physiological modalities and/or behavioural modalities [1], [2]. Commonly used physiological modalities include face, fingerprint, palm-print, ear, and iris, and behavioural modalities include gait, voice, signature, and key-stroke [3]. Compared with the traditional personal authentication systems with tokens or passwords, the biometric traits have the properties of universality, uniqueness, permanence and collectability [4], which makes the biometric systems more reliable. Among the various biometric traits, most of the previous works focus on face [5], fingerprint [6], and iris [7]. However, these traits suffer from several drawbacks such as, faces are changed through expressions and age, fingerprint readings are impacted by dirty or marked fingers, and iris reading is very expensive. Moreover, most of the physiological traits are active, and measurements for some traits may cause a worry to persons. Compared with these traits, ear has stable structure that is not changed through ages, and Iannarelli [4] has proven that there are no two ears having the same helix. Therefore, human ear satisfies the requirements of a biometric trait: uniqueness, universality, permanence and collectability.

Deep features extracted from the Convolutional Neural Networks (CNNs) have been demonstrated to obtain impressive performance in face verification [8] [9]. Compared with the handcrafted features, deep features have the advantages of discriminative and robust representation ability. The top layers contain more semantic information and describe the global feature of the images, while the intermediate layers describe the local features and the bottom layers contain more low-level information for the description of texture, edges et al. However, deep representation has not yet been applied for solving the ear recognition problem. Previous methods are mostly based on handcrafted features or their combination [10], [11], [12], [13], [14]. These methods are generally summarized as the following three types of ear features: 3D-based [10], appearance-based [15] and geometric features [13], [14]. For instance, the 3D-based features are more robust to illuminations and contain stable surface shape features, while the initial alignment may have large influence on matching results which leads to computational burden [12]. Appearance-based features are specifically designed by exploiting the color and texture properties of ear images [15], which may not handle various changes of images. Geometric features are easy to extract, and more robust to lighting conditions, which have shown to be effective for ear recognition problems [13]. In this work, we firstly exploit the deep features to represent ear images for ear recognition problem.

Matching process is a necessary process in ear recognition to generate scores indicating the similarity. While in the previous work, most similarity is measured based on various distances, such as Euclidean Distance [14], [16], Hamming Distance [17]. Recent years, discriminative classifiers, such as KNN and SVM are popularly adopted for classification in ear recognition task [15]. However, due to that the number of ear images per person is usually limited to $3 \sim 5$, SVM can not achieve the performance desired. Thus, we propose to apply pairwise SVM [18] to solve the problem. Pairwise SVM provides an extension of the binary SVM classifier to

978-1-5090-5507-4/17/$31.00 ©2017 IEEE
ICIS 2017, May 24-26, 2017, Wuhan, China

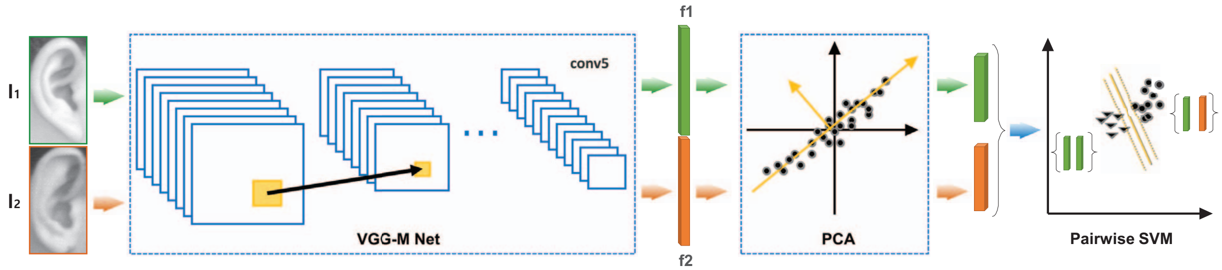

Fig. 1. Ear recognition framework based on the deep feature representation and Pairwise SVM.

multi-class classification, and which has been demonstrated to get excellent results on the classification problem.

The main contributions of the paper can be summarized as follows:

1) Deep features have more powerful representative ability compared with the existing handcrafted features, which are first utilized to represent the ear images for solving ear recognition.
2) Based on deep features, we firstly propose to use pairwise SVM for classification. And experimental results show its better recognition performance than most distance based matching methods and SVM.

The rest of the paper is organized as follows: Section 2 introduces the related work about the commonly used feature extraction methods and matching methods. Section 3 introduces the deep features extracted from VGG-M Net [19] and the pairwise SVM [18] for classification. Section 4 presents the experimental results and analysis. Finally, Section 5 draws the conclusion of the paper.

II. Related Work

A. Feature Extraction

Deep representations have been widely and successfully utilized in biometric systems, e.g. face recognition [20], spoofing attack detection [21]. However, the existing ear recognition methods are mainly based on the handcrafted features, which are mainly summarized as: 3D-based [10], [12], appearance-based [15] and geometric features [13], [14]. 3D-based features contain the structure information of the surface and are more robust to illuminations. However most 3D-based features need iterative closest point (ICP) approach for matching [22], which leads to the high computational cost and limits the wide use for civilian applications. Appearance-based features mainly aim to extract intensity, directional and spatial-temporal information holistically [23] or extract local ear features by exploiting the color and texture properties of ear images, such as color SIFT features and variants of local binary patterns (LBP) [15]. Ear shape is known to be quite complex and contains sufficient geometry information such as Ear Height Line (EHL), inner and outer helix, tragus and so on. Thus, shape based features have much less dimensions and are more robust to the noise and illumination variations. However, as the success of deep learning for feature engineering, it has been verified that, the top layer features have the property of describing the global feature of ear images, while the bottom layer features maintain more detailed information such as texture and edges, which shows that deep features can handle all those situations simultaneously.

B. Matching methods

In ear recognition problem, most previous works focus on the features extraction process, and kinds of holistic, local and hybrid approaches are proposed based on 2D and 3D features. For 2D-based features, researchers mainly adopt different distances as the matching criterion, such as Euclidean distance, and Hamming distance [14], [16], [17]. While on 3D-based features, ICP algorithm and its improvements are widely adopted to perform alignment and then the root mean square (RMS) registration error is used as the matching error criterion [12]. Recently, discriminative methods, such as KNN, SVM draw much attention [15]. In [24], SVM was firstly used for solving the ear recognition combined with PCA and ICA, which obtained better recognition effect. However, due to the lack of training images of ears and the multiple class property of ear database, SVM may not lead to the performance desired. Pairwise SVM extends the binary SVM to multi-class SVM from another perspective and has better interclass generalization performance, which can be applied to handle the multiple classes of ear image database better.

III. The Proposed Method

In this section, we present the deep features based pairwise SVM approach for ear recognition. Fig. 1 illustrates the framework of the proposed method, including the input pairwise images, the deep features extraction based on VGG-M Net [19], and the pairwise SVM classifier for decision.

A. CNN features extraction

In this paper, VGG-M Net is adopted to extract deep features for ear images due to the computational efficiency. VGG-M Net is trained on the ILSVRC dataset and it contains eight layers with weights, the first five are convolutional layers, and the remaining three are fully-connected layers. Specifically, the first and second convolutional layers are followed by its response normalization layers, and max-pooling layers follow both the response normalization layers as well as the fifth convolutional layer. Furthermore, the ReLU non-linearity is applied to the output of every convolutional and fully-connected layer. Table. I gives the VGG-M Net architectures used in this paper. Fig. 2 also visualizes the outputs from the convolutional layers corresponding to two sample images from the USTB I and II dataset separately.

We can see that these features do capture meaningful information about ears. Due to the property of CNN, we know that the top layer features maintain more semantic information to describe the global feature of ear images, while the middle and bottom layer features contain more local features, such as edges, textures. In experiments, we evaluate the convolutional layers, and the fully-connected layers separately to analyze the influence of different layer features for ear recognition. And due to the limitation of high dimension, we perform feature transformation by simply applying PCA [25].

TABLE I
THE VGG-M NET USED IN THIS PAPER.

Index of layer	Name	Type	Filter Size/Stride	Output Size
L1	conv1	Convolution	7×7 / 2	109× 109 × 96
L2	relu1	ReLU		109×109 × 96
L3	norm1	LRN		109×109 × 96
L4	pool1	Max Pooling	3×3 / 2	54× 54 × 96
L5	conv2	Convolution	5×5 / 2	26×26 × 256
L6	relu2	ReLU		26×26 × 256
L7	norm2	LRN		26× 26 × 256
L8	pool2	Max Pooling	3× 3 / 2	13× 13 × 256
L9	conv3	Convolution	3×3 / 1	13× 13 × 512
L10	relu3	ReLU		13× 13 × 512
L11	conv4	Convolution	3×3 / 1	13× 13 × 512
L12	relu4	ReLU		13× 13 × 512
L13	conv5	Convolution	3× 3 / 1	13× 13 × 512
L14	relu5	ReLU		13× 13 × 512
L15	pool5	Max Pooling	3×3 / 2	6 × 6 × 512
L16	fc6	fully connection		4096 × 1
L17	relu6	ReLU		4096 × 1
L18	fc7	fully connection		4096 × 1
L19	relu7	ReLU		4096 × 1
L20	fc8	fully connection		1000 × 1

B. Pairwise SVM

Pairwise SVM [18] relies on two input samples and predicts whether the two input samples belong to the same person or not. Due to the lack of samples of ear images, we select two images per person for training and the rest images for testing. Supposing that we have $\mathbf{n}$ person and denote the training set by $\mathbf{S} = \{\mathbf{x}_1^1, \mathbf{x}_1^2, \ldots, \mathbf{x}_n^1, \mathbf{x}_n^2\}$, where $\mathbf{x}_i^j$ denotes the j-th image of the i-th person. Based on the training set $\mathbf{S}$, we compound to pair samples $(\mathbf{x}_i^j, \mathbf{x}_p^q)$ and define $y_{ip} = +1$ if $i = p$ which we call the pair a positive pair, otherwise, $y_{ip} = -1$ and we call it a negative pair. The predict function obtained by pairwise SVM should be:

$$f\left(\mathbf{x}_i^j, \mathbf{x}_p^q\right) = \sum_{(m,n)\in \mathbf{I}} \alpha_{mn} y_{mn} K\left(\left(\mathbf{x}_m, \mathbf{x}_n\right), \left(\mathbf{x}_i^j, \mathbf{x}_p^q\right)\right) + b \tag{1}$$

where $\mathbf{I} \subseteq N \times N$ with $N = 2n$. According to [18], both balanced kernels or symmetric training sets can enforce the symmetry of a pairwise decision function equivalently. Due to the less ear images in ear recognition problem, we adopt the symmetric training sets strategy to complete pairwise SVM. Additionally, pairwise kernel:

$$K\left(\left(a, b\right), \left(c, d\right)\right) := \; k\left(a, c\right) + \; k\left(b, d\right) \tag{2}$$

is utilized as presented in [18]. Due to the high dimension of deep feature, even though PCA is performed, we also chose the linear kernel for all experiments.

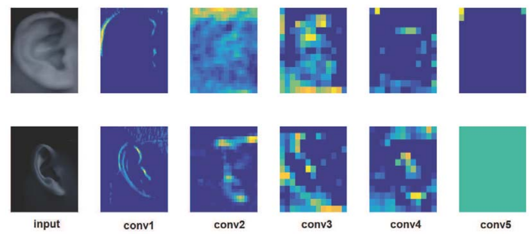

Fig. 2. The visualization of different deep features.

IV. EXPERIMENTAL RESULTS

A. Datasets and evaluation metrics

The experiments are performed on publicly available ear databases: USTB I and USTB II [26]. USTB databases are collected from the students and teachers at the department of information engineering of University of Science and Technology Beijing (USTB) [26]. In the following, we present a brief summary on each ear database. The USTB I database consists of 180 images from 60 subjects. Each subject has three images, including a normal frontal image, a frontal image with trivial rotation and an image with different illumination. All images have the same size 150 × 80 pixels. The USTB II database contains 308 images for 77 subjects, which is taken under illumination and angle change condition. Each subject has 4 images: the first image is the frontal ear image under standard illumination, the second and the third images with rotation are $+30$ and -30 respectively ear images and the fourth image is taken under weak illumination. All images on USTB II database have been acquired during Nov. 2003 to Jan. 2004 and have the same resolutions 400 × 300 pixel. Fig. 3 shows the original ear images provided by USTB I and USTB II, respectively.

For fair comparison with the sate-of-the-art approaches, we

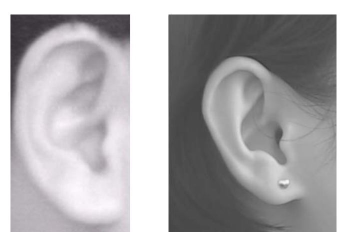

Fig. 3. Original ear images from USTB I and USTB II, respectively.

consider the commonly used scenario in experiments. That is, two images per person are utilized for training, and the left one image is used for testing for all databases. Finally the mean results of three random groups of experiments are constructed for robustness. We compare our method with the state-of-the-art methods [15], [16], [27]. Due to the high dimension of deep feature, PCA is performed before classification. We adopt kernel PCA technique for simplicity, and the dimension is fixed as the number of training samples. In the experiments, the rank-1 identification rate (i.e., recognition rate) is used as performance indicator. In evaluating the performance of pairwise SVM classifier, the Receiver Operating Characteristic (ROC) curve and the Equal Error Rate (EER) are used. ROC curve presents the relation between the true positive rate (TPR) and false positive rate (FPR), while the EER corresponds to the value when the false reject rate (i.e., 1-TPR) and false positive rate are equal along with the ROC curve.

B. Comparison with the state-of-the-arts

To evaluate the proposed approach, we construct experiments on the USTB I and USTB II databases compared with the existing state-of-the-arts [15], [16], [27]. Table. II presents the experimental results by using two images of each person for training. One can see that pairwise SVM performs better than most of them on both USTB I and II with the recognition rate of 98.3%, 92.04%. Among the compared methods, Benzaoui *et al.* proposed the local texture descriptors for ear image representation in [15], which includes local binary patterns, local phase quantization, and binarized statistical image features, furthermore three classifiers were adopted for recognition, such as KNN with city-block distance, SVM with linear kernel and SVM with RBF kernel. In [15], among the SVM based methods, the highest recognition rate of 97.4% is obtained on USTB I. However by using pairwise SVM with linear kernel, we get 98.3% recognition rate. A little higher recognition rate of 98.4% is obtained by Benzaoui *et al.* with the K-NN classifier using city-block distance. Actually, the recognition rate of the proposed method (98.3%) is comparable with that of Benzaoui's method (98.4%). In addition, pairwise SVM needs to compose positive and negative pair samples as input, which will lead to a vast amount of combinations. However in experiments, we randomly chose part of the total combinations which may lead to small variants in recognition rate. In summary, we can get comparable results with [15] on USTB dataset with 2 images for training. Since USTB II contains four ear images per person, thus we also perform experiments by using three images for training and the remaining one for testing. The results are presented in Table. III. It is obvious that pairwise SVM outperforms the traditional SVM and other state-of-the-art ear recognition methods with the recognition rate of 98.7%.

C. Performance Analysis

1) CNN layers: In order to analyze the representation ability of different layer features, we construct experiments on convolutional layer features and fully-connected layer features based on VGG-M Net. For simplicity, we chose $conv = \{L2, L6, L10, L12, L14\}$ as convolutional layer features and $fc = \{L16, L18, L20\}$ as fully-connected layer features.

TABLE II
RECOGNITION RATE (%) BY USING TWO IMAGE OF EACH PERSON FOR TRAINING ON USTB I AND USTB II.

Methods		USTB I (%)	USTB II (%)
Our approach	SVM	97.3	86.0
	Pairwise SVM	98.3	92.0
Benzaoui et al., 2014 [15]		98.4	−−
Ghoualmi et al. 2016 [16]		97.2	−−
Yaqubi et al. 2008 [28]		75.0	−−
Zhang et al. 2005 [29]	PCA	85.0	−−
	ICA	88.3	−−
Zhang and Mu 2008 [24]	PCA	85.0	−−
	ICA	91.7	−−
Nosrati et al. 2007 [30]	PCA	−−	84.3
	ICA	−−	87.7
	2D wavelet+PCA	−−	90.5

TABLE III
RECOGNITION RATE (%) BY USING THREE IMAGES OF EACH PERSON FOR TRAINING ON USTB II.

Methods		USTB II (%)
Our approach	SVM	97.5
	Pairwise SVM	98.7
Tariq et al. 2011 [27]		96.1
Guo and Xu, 2008 [31]		93.2
Ghoualmi et al. 2016 [16]	HE	93.5
	CLAHE	94.2
	ABC	94.8
Zhang and Mu, 2008 [24]	PCA	81.8
	ICA	92.2

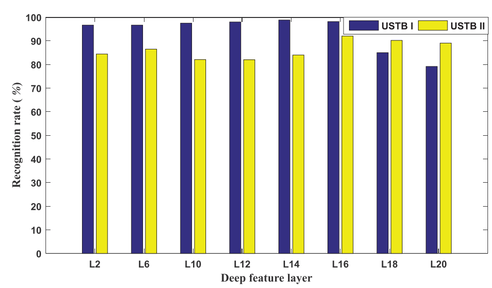

Fig. 4. The recognition rate (%) of different convolutional layer and fully-connected features on USTB I and II.

As shown in Fig. 4. In the USTB I experiments, the convolutional layer features lead to comparable performance compared with the fully-connected layer features with $rank1 = \{96.67, 96.67, 97.50, 98.00, 98.30, 98.17, 85.00, 79.17\}(\%)$. However, for the USTB II database, the fully-connected layer features perform better than the bottom and middle layers with the recognition rate of $\{84.42, 86.49, 82.08, 82.00, 84.01, 92.04, 90.20, 89.05\}(\%)$. The results can be explained by that, the ear images in USTB I contains the complete and clear ear without background information, while that in USTB II not only contains the ear, but also part of the area around the ear, which introduces noise, as Fig. 3 shows. Since that top features of CNN contain more semantic information and the middle and bottom features contain more low-level information, which explains that the low-level information improves the representation ability of ear images in USTB I, while the semantic information on ”ear” plays more important role in USTB II for ear recognition.

2) Model parameters: For comprehensive analysis, we also evaluate the performance of the trade-off C with the output of last convolutional feature on USTB I. Fig. 5 presents the recognition rates of different values of trade-off C. One can see that, the optimal value of C locates in the range of $[1, 100]$ with the recognition rate around 98.3%. Smaller or larger values of C may lead to under-fitting and over-fitting, respectively. In experiments, we set C a default value with 1.

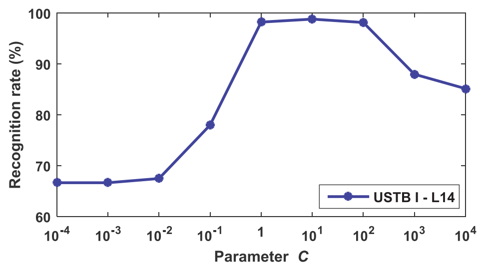

Fig. 5. The analysis on trade-off parameter C.

3) Classifiers: Due to the lack of ear images per person, SVM can not achieve the performance desired. Thus, we construct experiments to evaluate the performance of the proposed pairwise SVM and the standard SVM. Among experiments, the last convolutional feature is utilized for both pairwise SVM and SVM.

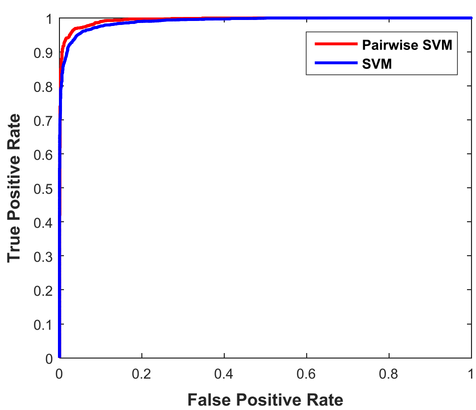

Fig. 6. The ROC curves of SVM and Pairwise SVM.

Fig. 6 shows the ROC curves of pairwise SVM and SVM on USTB I, and Fig. 7 shows their cropped one for clearness. From Figs. 6 and 7, we can see that, pairwise SVM outperforms SVM obviously which validates the effectiveness of pairwise SVM in ear recognition. For simplicity, we also draw a conclusion of ROC curves with the EER indictor. As shown in Table. IV, pairwise SVM gets a higher recognition rate of 98.3% with a lower EER 0.042, compared with the SVM with the recognition rate of 97.3% and the ERR of 0.056.

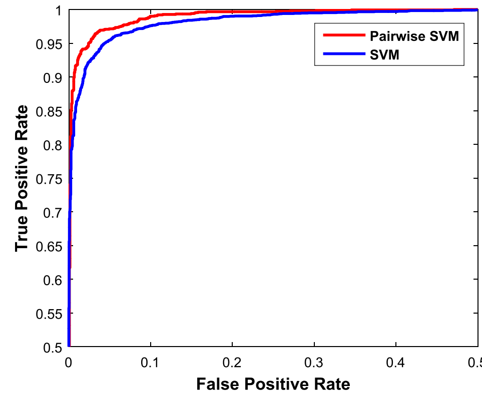

Fig. 7. The ROC curves crop of SVM and Pairwise SVM.

TABLE IV
COMPARISON OF RECOGNITION RATE AND EER ON USTB-1.

Methods	Recognition rate (%)	EER
Pairwise SVM	98.3	0.042
SVM	97.3	0.056

4) Computational efficiency: The contributions of this paper are mainly the application of deep features for representing the ear images and the utility of pairwise SVM for solving the classification problem in ear recognition task as described in Section 1. For the efficiency analysis, time consuming is from three parts in our proposed method, i.e. the deep feature extraction, the PCA process and the classification. The deep features can be extracted in advance separately. And the time cost of PCA operation depends on the feature dimension in different layers. For classification, we adopt the symmetric training sets strategy and linear kernel to complete pairwise SVM, which has similar time consuming with traditional SVM and depends on the feature dimension.

V. CONCLUSION

In this paper, we firstly propose to use deep features extracted from the Convolutional Neural Network (CNN) to represent the ear images for solving the ear recognition task. In order to deal with the lack of training ear images of each person, pairwise SVM is adopted to deal with the multi-class classification for ear recognition. Finally, we evaluate the deep features based pairwise SVM for ear recognition with the state-of-the-art approaches. The experimental results achieved a promising recognition rates and demonstrated the superiority of the proposed approach.

Acknowledgments

This work was supported in part by the co-operation between Higher Education Commission of Egypt and Chinese Government, the National Natural Science Foundation of China (Grant No.s: 61671182, 61471146 and 51308096), and the

Foundation of Education Department of Heilongjiang Province (Grant No.: 12541050).

REFERENCES

[1] S. Z. Li, *Encyclopedia of Biometrics: I-Z.* Springer Science & Business Media, 2009, vol. 1.

[2] C. Le and R. Jain, "A survey of biometrics security systems," *EEUU. Washington University in St. Louis*, 2009.

[3] A. Jain, P. Flynn, and A. A. Ross, *Handbook of biometrics*. Springer Science & Business Media, 2007.

[4] A. V. Iannarelli, *Ear identification*. Paramont Publishing Company, 1989.

[5] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[6] D. Maltoni, D. Maio, A. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

[7] J. Daugman, "New methods in iris recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 37, no. 5, pp. 1167–1175, 2007.

[8] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1891–1898.

[9] G. B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2518–2525.

[10] S. M. Islam, R. Davies, M. Bennamoun, and A. S. Mian, "Efficient detection and recognition of 3d ears," *International Journal of Computer Vision*, vol. 95, no. 1, pp. 52–73, 2011.

[11] K. Annapurani, M. Sadiq, and C. Malathy, "Fusion of shape of the ear and tragus–a unique feature extraction method for ear authentication system," *Expert Systems with Applications*, vol. 42, no. 1, pp. 649–656, 2015.

[12] H. Chen and B. Bhanu, "Human ear recognition in 3d," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 718–737, 2007.

[13] M. Rahman, M. S. Sadi, and M. R. Islam, "Human ear recognition using geometric features," in *Electrical Information and Communication Technology (EICT), 2013 International Conference on*. IEEE, 2014, pp. 1–4.

[14] I. Omara, F. Li, H. Zhang, and W. Zuo, "A novel geometric feature extraction method for ear recognition," *Expert Systems with Applications*, vol. 65, pp. 127–135, 2016.

[15] A. Benzaoui, A. Hadid, and A. Boukrouche, "Ear biometric recognition using local texture descriptors," *Journal of Electronic Imaging*, vol. 23, no. 5, pp. 053 008–053 008, 2014.

[16] L. Ghoualmi, A. Draa, and S. Chikhi, "An ear biometric system based on artificial bees and the scale invariant feature transform," *Expert Systems with Applications*, vol. 57, pp. 49–61, 2016.

[17] T.-S. Chan and A. Kumar, "Reliable ear identification using 2-d quadrature filters," *Pattern Recognition Letters*, vol. 33, no. 14, pp. 1870–1881, 2012.

[18] C. Brunner, A. Fischer, K. Luig, and T. Thies, "Pairwise support vector machines and their application to large scale problems," *Journal of Machine Learning Research*, vol. 13, no. Aug, pp. 2279–2292, 2012.

[19] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," *arXiv preprint arXiv:1405.3531*, 2014.

[20] M. M. Ghazi and H. K. Ekenel, "A comprehensive analysis of deep learning based representation for face recognition," *arXiv preprint arXiv:1606.02894*, 2016.

[21] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcao, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864–879, 2015.

[22] P. Yan and K. W. Bowyer, "Biometric recognition using 3d ear shape," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 8, pp. 1297–1308, 2007.

[23] D. Snchez and P. Melin, "Optimization of modular granular neural networks using hierarchical genetic algorithms for human recognition using the ear biometric measure," *Engineering Applications of Artificial Intelligence 27 (2014): 41-56*, 2014.

[24] H. Zhang and Z. Mu, "Compound structure classifier system for ear recognition," in *2008 IEEE International Conference on Automation and Logistics*. IEEE, 2008, pp. 2306–2309.

[25] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *International Conference on Artificial Neural Networks*. Springer, 1997, pp. 583–588.

[26] Z. Mu, "Ustb ear image database," 2005. [Online]. Available: http://www1.ustb.edu.cn/resb/en/index.htm

[27] A. Tariq, M. A. Anjum, and M. U. Akram, "Personal identification using computerized human ear recognition system," in *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*, vol. 1. IEEE, 2011, pp. 50–54.

[28] M. Yaqubi, K. Faez, and S. Motamed, "Ear recognition using features inspired by visual cortex and support vector machine technique," in *Computer and Communication Engineering, 2008. ICCCE 2008. International Conference on*. IEEE, 2008, pp. 533–537.

[29] H.-J. Zhang, Z.-C. Mu, W. Qu, L.-M. Liu, and C.-Y. Zhang, "A novel approach for ear recognition based on ica and rbf network," in *2005 International Conference on Machine Learning and Cybernetics*, vol. 7. IEEE, 2005, pp. 4511–4515.

[30] K. F. Nosrati, Masoud S. and F. Faradji, "Using 2d wavelet and principal component analysis for personal identification based on 2d ear structure," *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on. IEEE*, 2007.

[31] Y. Guo and Z. Xu, "Ear recognition using a new local matching approach," *15th IEEE International Conference on Image Processing. IEEE*, 2008.